Claims

We claim:

1. A method for monitoring data values in an event monitoring system, using a set of parameters characterizing a reference distribution of the data values, the method comprising the steps of:

computing from a given one of the data values a corresponding probability value, utilizing the set of parameters characterizing the reference distribution;

performing a thresholding operation on the probability value;

generating alarm information based on the result of the thresholding operation; and

periodically updating the set of parameters to take into account one or more of the data values.

2. The method of claim 1 wherein the event monitoring system comprises an event counter which generates the data values and an event monitor which processes the data values to generate the alarm information.

3. The method of claim 1 wherein the set of parameters comprises at least a mean and a variance for the reference distribution of the data values.

4. The method of claim 3 wherein the reference distribution comprises an asymmetric distribution.

5. The method of claim 4 wherein the asymmetric distribution comprises one of a negative binomial distribution, a lognormal distribution and a Poisson distribution.

6. The method of claim 1 wherein the data values are collected and thresholded in accordance with a measurement timescale, the set of parameters are updated in accordance with a parameter timescale longer in duration than the measurement timescale, and a complete set of the

parameters is determinable in accordance with a cycle timescale longer in duration than the parameter timescale.

7. The method of claim 6 wherein the measurement timescale comprises approximately one minute, the parameter timescale comprises approximately one hour and the cycle timescale comprises approximately one day.

8. The method of claim 1 wherein the probability value computed from a given data value x comprises at least one of:

(i) an upper tail probability $P[data\ value \geq x]$;

(ii) a lower tail probability $P[data\ value \leq x]$; and

(iii) a minimum of $\{P[data\ value \geq x], P[data\ value \leq x]\}$;

wherein the probability value is computed using the set of parameters characterizing the reference distribution that applies at the time the given data value x is collected.

9. The method of claim 1 wherein the set of parameters characterizing the reference distribution is determined at least in part by interpolating coefficients over multiple parameter timescales.

10. The method of claim 1 further comprising the step of periodically monitoring validity of the reference distribution.

11. The method of claim 10 wherein the validity of the reference distribution is monitored by generating a histogram of a plurality of probability values and determining if counts associated with a plurality of intervals of the histogram are each approximately the same.

12. The method of claim 11 wherein the probability values utilized in generating the histogram are corrected for continuity by:

computing the probability of observing data value x under the reference distribution as $p_x = p_L + p_U - 1$, where p_L and p_U denote respective lower and upper tail probabilities associated with data value x under the reference distribution;

taking a random draw Z from a uniform distribution on [0, 1]; and

computing a continuity corrected probability value as $p_{cont} = p_L - Z * p_x$.

13. The method of claim 1 further comprising the step of determining if the given data value is an outlier, and if so modifying the given data value as follows:

replacing an upper tail outlier with a random draw from the reference distribution conditioned to be within a specified upper portion of the distribution, such that the upper tail outlier is replaced with another upper tail data value; and

replacing a lower tail outlier with a random draw from the reference distribution conditioned to be within a specified lower portion of the distribution, such that the lower tail outlier is replaced with another lower tail data value.

14. The method of claim 1 wherein the step of updating the set of parameters characterizing the reference distribution further comprises updating a given one of the parameters as a weighted average of corresponding parameter values determined over a designated timescale.

15. The method of claim 1 further including the step of initializing the set of parameters characterizing the reference distribution.

16. The method of claim 1 wherein the alarm information further comprises an alarm severity measure that indicates the severity of alarm conditions based on multiple-mode error comprising at least error duration and error spread.

17. The method of claim 1 wherein the alarm information further comprises an alarm severity measure generated at least in part utilizing at least one of a ratio of multiple probability values and a difference of logarithms of probability values.

18. The method of claim 1 wherein the alarm information further comprises an alarm severity measure S, and the given data value has a logit transformation L, for a corresponding probability value p, given by:

$$L(p) = log\left(\frac{1-p}{p}\right),$$

and further wherein the alarm severity measure is updated as follows:

$$S^{new} = (1-w)*S + w*L,$$

where w denotes a severity weight.

19. An apparatus for monitoring data values in an event monitoring system, using a set of parameters characterizing a reference distribution of the data values, the apparatus comprising:

a memory for storing the set of parameters; and

a processor coupled to the memory and operative to control operations associated with the monitoring of the data values, the operations including:

computing from a given one of the data values a corresponding probability value, utilizing the set of parameters characterizing the reference distribution;

performing a thresholding operation on the probability value;

generating alarm information based on the result of the thresholding operation; and

periodically updating the set of parameters to take into account one or more of the data values.

20. An article of manufacture comprising a machine-readable storage medium for storing one or more programs for use in monitoring data values in an event monitoring system, using a set of parameters characterizing a reference distribution of the data values, wherein the one or more programs when executed implement the steps of:

computing from a given one of the data values a corresponding probability value, utilizing the set of parameters characterizing the reference distribution;

performing a thresholding operation on the probability value;

generating alarm information based on the result of the thresholding operation; and

periodically updating the set of parameters to take into account one or more of the data values.